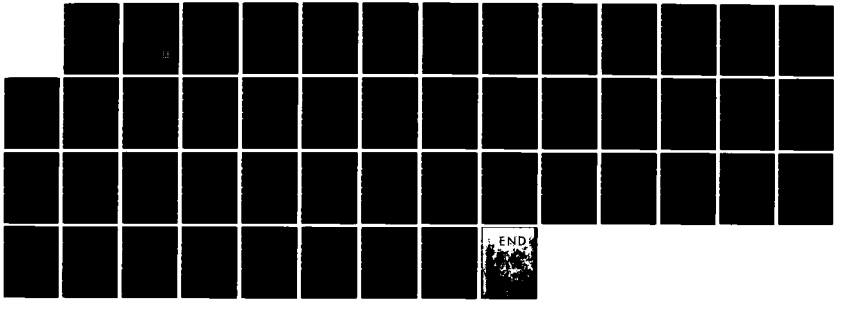
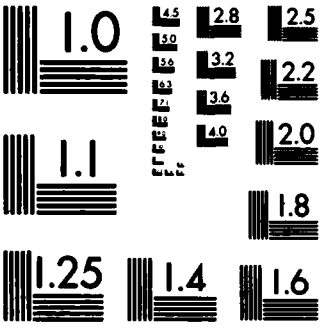
AD-A130 213 ON PERTURBATION THEORY(U) MASSACHUSETTS INST OF TECH 1/1
LEXINGTON LINCOLN LAB L G TAFF 01 JUN 83 TR-650
ESD-TR-83-023 F19628-80-C-0002

UNCLASSIFIED F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ESD-TR-83-023

(12)

AD A130 043

Technical Report 650

L.G. Taff

On Perturbation Theory

1 June 1983

Prepared for the Department of the Air Force
under Electronic Systems Division Contract F19628-80-C-0002 by

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LEXINGTON, MASSACHUSETTS

DTIC FILE COPY

Approved for public release; distribution unlimited.

DTIC
JUL 11 1983

83 07 08 020

A

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the support of the Department of the Air Force under Contract F19628-80-C-0002.

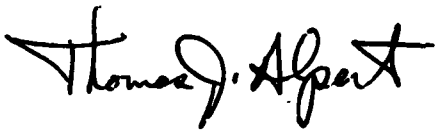
This report may be reproduced to satisfy needs of U.S. Government agencies.

The views and conclusions contained in this document are those of the contractor and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the United States Government.

The Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER



Thomas J. Alpert, Major, USAF
Chief, ESD Lincoln Laboratory Project Office

Non-Lincoln Recipients

PLEASE DO NOT RETURN

Permission is given to destroy this document when it is no longer needed.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

ON PERTURBATION THEORY

L.G. TAFF
Group 94

TECHNICAL REPORT 650

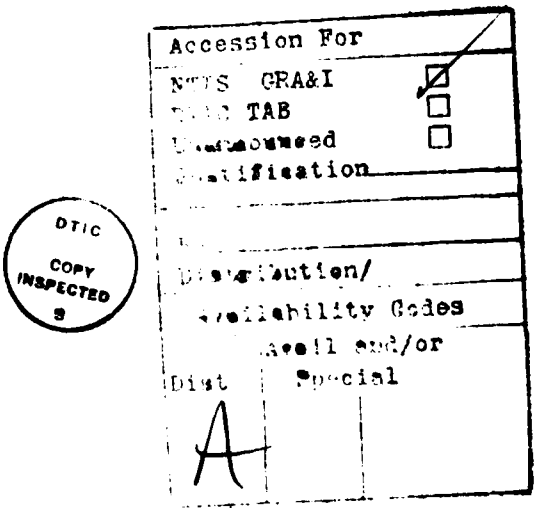
1 JUNE 1983

Approved for public release; distribution unlimited.

LEXINGTON **MASSACHUSETTS**

ABSTRACT

Perturbation theory in celestial mechanics has a long, rich history of failure stretching back to Newton. I believe that the causes of this are two-fold. One problem is the difficulty of dealing with the mathematical structure used in celestial mechanics to express perturbation theory as opposed to the constructs used in field theories (eg. trajectory equations vs. linear second order partial differential equations). The second flows directly from this and relates to the misapplications of certain mathematical techniques (averaging, series expansions) within the context of perturbation theory. These incorrect analyses usually appear in second-order theories such as Kozai's (1959) artificial satellite theory. Ideally this report would clearly illustrate the nature of these difficulties utilizing a complex (but exactly soluble) physical model intimately tied to the two-body problem and then go on to lay the foundations for a new perturbation theory. I believe that that's exactly what is accomplished herein except that the hints of the base of this new mathematical formalism are severely limited. The exactly soluble physical model is the three dimensional harmonic oscillator complicated by anisotropic terms, anharmonic terms, and air resistance. The deep connection is provided by Bertrand's theorem which is also proved.



CONTENTS

	ABSTRACT	iii
I.	INTRODUCTION	1
II.	BERTRAND'S THEOREM	4
III.	THE ISOTROPIC THREE DIMENSIONAL, SIMPLE HARMONIC OSCILLATOR	9
	A. The Orbit	10
	B. The Time Dependence	12
	C. A Perturbation Theory	14
IV.	THE ANISOTROPIC, THREE DIMENSIONAL, SIMPLE HARMONIC OSCILLATOR	17
	A. The Mathematical Foundations of Perturbation Theory	17
	B. The Method of Averaging	25
	C. A New Representation	30
V.	THE ANHARMONIC, ONE DIMENSIONAL, HARMONIC OSCILLATOR	36
VI.	CONCLUDING REMARKS	40
	References	42

I. INTRODUCTION

There is no problem in celestial mechanics that perturbation theory provides a satisfactory solution to (of which I am aware). Saturn's rings, the Kirkwood gaps, Jupiter's satellites, the Sun-Jupiter-Saturn system, etc., all exhibit resonances, unusual structures, or some other peculiarity not yet deduced from celestial mechanics. Celestial mechanics may provide explanations (as in the fact that the Kirkwood gaps occur at simple multiples of Jupiter's mean motion) but not predictions. Why is this state of affairs so poor? I believe that the reason is two-fold (at least). First of all celestial mechanics is a trajectory theory and not a field theory. The field theories with good perturbation techniques rely on linear second order differential equations not non-linear first order ones (eg. the Hamilton-Jacobi equation). Secondly the mathematical nature of the problem in celestial mechanics is compounded by a poor choice of coordinate systems. It does no good to express perturbation theory in terms of the orbital element set, because we will then immediately and intuitively understand the result, if we can never solve any problems in this coordinate system. This becomes compounded when our frustration with our failures further leads us astray in the use of unjustifiable analytic techniques. Examples of the latter include incorrect power series expansions and the method of averaging.

It is possible that the reader is not as aware of the shortcomings and pitfalls of the misapplications of perturbation theory in celestial mechanics as the above paragraph suggests is the case. This report will serve to inform the reader of the situation by a carefully constructed sequence of illustrations. As the nature of the demonstration is by counter-example, it follows

that each counter-example must in itself be exactly soluble. Moreover it is clear that no such meaningful example is likely to come from the normal realm of problems dealt with in celestial mechanics. However it would be preferable to deal with a system as close to the two-body problem as possible. The existence of Bertrand's (1873) theorem is used to find the suitable connection and the alternative physical model--the three dimensional, isotropic, simple harmonic oscillator.

Bertrand's theorem simply states that of all possible central force potential functions only two give rise to bounded, closed orbits. One is the two-body potential ($-k/r$) and the other is the isotropic, simple harmonic oscillator ($kr^2/2$). The next section follows Arnold's (1978) demonstration of this result. Following that, motion under the three dimensional, isotropic simple, harmonic oscillator potential is discussed (the orbits are ellipses with force center at the center). A perturbation theory for these orbits is developed and a set of first order differential equations is derived. These are exactly analogous to Lagrange's variation of parameter equations in the two-body problem.

The fourth Section deals with the mathematical foundations of perturbation and an attempt to use the variation of parameters equations when a perturbation yielding anisotropy is applied. This attempt very closely follows Kozai's (1959) artificial satellite theory in format. The purpose is not to criticize Kozai but rather to illustrate the invalidity of a technique widely used in celestial mechanics. This section concludes with the development of a new set of variation of parameters equations utilizing a different set of constants (that is other than the orbital element set).

In this representation the perturbation equations are exactly soluble to all orders--indeed the whole system is exactly soluble in closed form. The misleading inferences of the classical approach can be clearly seen now.

The fifth section introduces anharmonic perturbations. These too, in the right basis, present a set of perturbation equations for which we can obtain the solution to all orders (explicitly) as well as in closed form. This further elucidates the nature of the failings of other approaches.

In the final section I discuss non-conservative perturbations due to drag and "third-body" forces for the harmonic oscillator. This is very short as are some remarks concerning additional representations. By now I've proved my points and our energies should be devoted to productive rearrangements of perturbation theory for the two-body problem.

II. BERTRAND'S THEOREM

According to Plummer (1908), in 1873 Bertrand proved that the inverse square law force and the direct linear force law are the only two central force laws, derivable from a potential function, which yield bounded closed orbits. On pages 2-9 of his book Plummer presents a "proof" of this result. Goldstein's second edition of Classical Mechanics (1982) reproduces Plummer's result in modern language. It does not appear to me that this proof is rigorous. Arnold (1978), without ever mentioning Bertrand, does provide the outlines of a rigorous proof of this result. While I shan't fill in all of the analytical steps for the reader, immediately below I flesh out Arnold's logic. The importance of the result for this discussion is that it yields a deep connection between the two-body problem and the three dimensional, isotropic, simple harmonic oscillator problem. (I felt it incumbent upon me to debunk the misapplications of perturbation theory via a physical situation either identical to the $1/r$ potential or intimately associated with it.)

The logic of Arnold's argument can be simply outlined. First of all, if all bounded orbits are closed, then since a circular orbit is always possible for a central force, all nearly circular bounded orbits should be closed too. Moreover this (eg. being closed) must be true independent of the radius of the circular orbit. The implications of these statements are that the potential energy is either a power law $U(r) = Ar^p$ ($p \geq -2, p \neq 0$) or a logarithmic function $U(r) = B\ell nr$ (the $p=0$ case). By considering the constraints of "closedness" on the apsidal angle Arnold then shows that the logarithmic case can not be closed while there are two possibilities for the power law case. These two deal with positive power laws ($p>0$) and the

meaning of bounded and with negative power laws which yield infinitely tight binding. After an examination of these two limiting cases it turns out that only p=+2 and p=-1 are left. QED.

The conservation of energy equation for a central force after the usual reduction to the equivalent one-body problem (cf. Eq. 2 below) yields

$$\frac{m\dot{r}^2}{2} + \frac{L^2}{2mr^2} + U = E$$

Hence the "pseudo-potential energy" $V = U + L^2/2mr^2$. From this we have that

$$\int dt = \pm \int dr \; / \; [2(E-V)/m]^{1/2}$$

and from $L = mr^2\dot{\theta}$,

$$\theta = \int \frac{(L/r^2)dr}{[2m(E-V)]^{1/2}}$$

Orbits exist such that $E \geq V(r)$ for a fixed value of L. This inequality gives one (or several depending upon V) regions with

$$0 \leq r_{min} \leq r \leq r_{max} \leq \infty$$

Clearly if $0 \leq r_{min} < r_{max} < \infty$ then the motion is bounded ($r_{max} < \infty$) and also takes place in the annulus with inner radius r_{min} and outer radius r_{max}. As θ increases r varies between these two limits. The angular extent between successive apsides (extrema of r) is called the apsidal angle. It is given by the expression

$$\Theta = \int_{r_{min}}^{r_{max}} \frac{(L/r^2)dr}{\{2m[E-V(r)]\}^{1/2}} \qquad (1)$$

A closed orbit is one such that Θ is commensurate with 2π. Finally in the special case when E is equal to a minimum of V then $r_{min} = r_{max} = R$ and the orbit is a circle of radius R. Now that the stage has been set, the argument can be filled in.

First let us calculate the value of Θ for a nearly circular orbit. Since the orbit is nearly circular E must be slightly larger than the minimum value of V. Near r=R V can be written as (since it's a minimum V'(R) = 0, V"(R)>0)

$$V(r) = V(R) + \frac{(r-R)^2}{2} V''(R)$$

The extremes of r are found by solving E-V(r) = 0. They are

$$R \pm \{2[E-V(R)]/V''(R)\}^{1/2}$$

with r_{max} given by the plus sign and r_{min} given by the minus sign. Use the definition in Eq. (1) of Θ, plus the change of variables

$$x = (r-R)\left[\frac{V''(R)}{2[E-V(R)]}\right]^{1/2}$$

to get

$$\Theta = L/[V''(R)]^{1/2} \int_{-1}^{1} \frac{dx}{(1-x^2)^{1/2}} \left[R + x\left\{\frac{2[E-V(R)]}{V''(R)}\right\}^{1/2}\right]^{-2}$$

so

$$\Theta_{circ} \simeq \pi L/R^2\sqrt{V''(R)m}$$

In terms of U,

$$\Theta_{circ} \simeq \pi[U'/(3U'+RU'')]^{1/2}/\sqrt{m}$$

Now impose the fact that we want this result to be independent of the value of R. This simply means that

$$\frac{U'}{3U'+RU''} = \text{constant}$$

Choose the form of the constant to be 1/(2+p) and integrate. One gets directly that

$$U = Ar^p \quad (p \geq -2, p \neq 0) \text{ or } U = B\ell nr$$

within an arbitrary, unimportant, additive constant. From the formula for Θ_{circ} we now get that (set m=1)

$$\Theta_{circ} = \pi/(2+p)^{1/2} \qquad (p=0 \text{ is the logarithmic case})$$

This completes the first part of the result.

If p=0 (the logarithmic case) then $\Theta_{circ} = \pi/\sqrt{2}$ which is not a rational multiple of 2π. Hence this functional form can be dismissed from further consideration. If p>0 then as $r \to \infty$, $U(r) \to \infty$. Therefore, to remain bounded, $E \to \infty$ too. The <u>general</u> limit as $E \to \infty$ of Θ is $\pi/2$. To see this make the change of u=L/r in Eq. (1). Then

$$\Theta = \int_{u_{min}}^{u_{max}} \frac{dx}{\{2m[E-W(x)]\}^{1/2}}$$

where $W(x) = U(L/x) + x^2/2m$. Further let $y = x/u_{max}$ so

$$\Theta = \int_{y_{min}}^{1} dy/\{2m[w(1)-w(y)]\}^{1/2}$$

where $w(y) = y^2/2m + U(L/y\ u_{max})/u^2_{max}$. Clearly as $E \to \infty$, $u_{max} \to \infty$, $y_{min} \to 0$ and the last term in w is negligible. The $\pi/2$ result follows after a trivial integration. Now if $\pi/(2+p)^{1/2} = \pi/2$ for $p>0$, $p=2$ and we have the isotropic harmonic oscillator as a possibility.

The last piece involves $p<0$. Here there's the possibility that $E \to -\infty$ as $r \to 0$. First we show that in general as $E \to -\infty$, $\Theta = \pi/(2-q)$ for power law potential energies $U = -kr^{-q}$, $q \in (0,2)$. The proof is as above and leads to

$$\Theta = \int_{0}^{1} dx/(x^q - x^2)^{1/2} = \pi/(2-q)$$

(This is one of several misprints in this section of Arnold's book.) Now replace -q by p and demand equivalence with the circular orbit result or that

$$\pi/(2+p) = \pi/(2+p)^{1/2}$$

Only $p = -1$ satisfies this case--the inverse square law.

III. THE ISOTROPIC, THREE DIMENSIONAL, SIMPLE HARMONIC OSCILLATOR

Just to be sure that we understand the terminology the potential energy is $U = kr^2/2$. The force constant is k and it is the same in all directions (isotropic). If the force constant differed in different directions, then $U = \underline{r}\cdot k\cdot\underline{r}/2$ where k is now a diagonal tensor

$$k = \begin{pmatrix} k_x & 0 & 0 \\ 0 & k_y & 0 \\ 0 & 0 & k_z \end{pmatrix}$$

This would be an anisotropic, three dimensional, simple harmonic oscillator. Since $r^2=x^2+y^2+z^2$ this explains the three dimensional aspect. The simple harmonic part is due to the fact that U is quadratic in x,y,z. The resulting motion is then sinusoidal with a period independent of the amplitude. The next Section discusses anisotropic oscillators and the one following that anharmonic ones.

The force is $\underline{F} = -\nabla U = -k\underline{r} = m\ddot{\underline{r}}$. Energy conservation is obvious and we set

$$E = U + \frac{m}{2}\dot{\underline{r}}\cdot\dot{\underline{r}} \geq 0 \qquad (2)$$

Furthermore since $\underline{F}$ is central $\underline{L} = \underline{r} \times m\dot{\underline{r}}$ is a constant vector. Therefore the orbit lies in a plane given by $\underline{r}\cdot\underline{L} = 0$ and the problem can be reduced to a two dimensional one. Set

$$\underline{L} = L\,(\sin i \sin\Omega, -\sin i \cos\Omega, \cos i) \qquad (3)$$

where i (the inclination) and Ω (the longitude of the ascending node) specify the orientation of $\underline{L}$ and L is its magnitude. Finally introduce polar coordinates r,θ in the orbital plane so that the equations of motion are

$$m(\ddot{r} - r\dot{\theta}^2) = f(r) = -dU/dr = -kr$$

$$\frac{m}{r}\,\frac{d}{dt}\,(r^2\dot{\theta}) = 0$$

A. The Orbit

As usual introduce Binet's transformation $u = 1/r$ and switch from time derivatives to derivatives with respect to θ via (prime = $d/d\theta$)

$$\dot{r} = r'\dot{\theta} = -u'\dot{\theta}/u^2 = -Lu'/m$$

Then we find that the equations of motion are

$$u'' + u = \frac{-mf(u^{-1})}{L^2u^2} = \frac{mk}{L^2u^3} \tag{4}$$

Define the mean motion n by

$$n^2 = k/m \tag{5}$$

A single integration of Eq. (4) gives

$$(u')^2 = \text{constant} - \left(u^2 + \frac{m^2n^2}{L^2u^2}\right)$$

Comparison with the conservation of energy equation (2) in the u form shows that the constant is $2Em/L^2$. Hence,

$$(u')^2 = \frac{(u_-^2 - u^2)(u^2 - u_+^2)}{u^2}$$

where

$$u_+^2 = 1/a^2,\ u_-^2 = 1/b^2 = 1/a^2(1-e^2)$$

and

$$r_{max}^2 = a^2 = \frac{L^2/m}{E-(E^2-L^2n^2)^{1/2}} = \frac{E+(E^2-L^2n^2)^{1/2}}{mn^2} \tag{6a}$$

$$r_{min}^2 = b^2 = a^2(1-e^2) = \frac{L^2/m}{E+(E^2-L^2n^2)^{1/2}} = \frac{E-(E^2-L^2n^2)^{1/2}}{mn^2} \tag{6b}$$

At $r = a$ or b $\dot{r} = 0$ and $r' = 0$ hence their identification with the extrema of r.

Suppose that the motion starts at $\theta = \theta_0$ where $r = a$. Then thereafter $r' < 0$ so

$$\int_{1/a}^{u} \frac{x\,dx}{[(b^{-2}-x^2)(x^2-a^{-2})]^{1/2}} = +\int_{\theta_0}^{\theta} d\phi$$

or

$$r^2 = \frac{a^2(1-e^2)}{1-e^2+e^2\sin^2(\theta-\theta_0)}$$

Without loss of generality choose $\theta_0=0$ and let another orientation angle, ω, define the line of the major axis ("the argument of periapse"). Then

$$r^2 = \frac{a^2(1-e^2)}{1-e^2+e^2\sin^2\theta} \tag{7}$$

is the equation of the orbit. If $\xi = r\cos\theta$, $\eta = r\sin\theta$ are rectangular

coordinates in the orbital plane, then this is equivalent to

$$\xi^2/a^2 + \eta^2/b^2 = 1$$

This is the equation of an ellipse, center at the origin, major axis a, minor axis b (and therefore eccentricy e), whose major axis (the line of apsides) lies on $\eta = 0$.

B. The Time Dependence

Return to the conservation of energy equation (2) and solve for $\dot{r}^2$,

$$\dot{r}^2 = \frac{n^2}{r^2}(a^2-r^2)(r^2-b^2)$$

Suppose r=a, θ=0 corresponds to t=T ("the time of periapse passage"). Since $a \geq r$, $\dot{r} < 0$ for t>T so

$$\int_a^r \frac{x\,dx}{[(a^2-x^2)(x^2-b^2)]^{1/2}} = -n\int_T^t d\tau$$

or

$$r^2 = a^2\cos^2 M + b^2 \sin^2 M \qquad (8)$$

where the mean anomaly M=n(t-T). Note that θ is the true anomaly and the "eccentric anomaly" too.

However $r^2 = \xi^2+\eta^2$ so

$$\xi = a\cos M,\ \eta = b\sin M \qquad (9)$$

Note too that $L = mr^2\dot{\theta} = 2mdA/dt$ where A is the areal rate. Whence,

$$L = 2m(\pi ab/P) = mnab \tag{10}$$

where the period is $P = 2\pi/n$.

To determine the time dependence of θ use $L=mr^2\dot{\theta}$ and the expression (8) for $r^2(t)$, viz.

$$\int_0^\theta d\phi = \frac{L}{m}\int_T^t \frac{d\tau}{a^2\cos^2\mu+b^2\sin^2\mu} \quad , \mu=n(\tau-T)$$

The result is

$$\tan\theta = (1-e^2)^{1/2}\tan M, \; M=n(t-T) \tag{11}$$

This is "Kepler's equation".

To summarize the three dimensional motion is given by

$$\underline{r} = S\underline{q} = rS\begin{pmatrix}\sin\theta\\ \cos\theta\\ 0\end{pmatrix} = aS\begin{pmatrix}\cos M\\ (1-e^2)^{1/2}\sin M\\ 0\end{pmatrix} \tag{12}$$

where the unitary rotation matrix $S=R_3(-\Omega)R_1(-i)R_3(-\omega)$. The R matrices are the usual elementary rotation matrices. The angular momentum has already been given; the energy E is

$$E = \frac{mn^2a^2}{2}\,(2-e^2) \tag{13}$$

C. A Perturbation Theory

Now, because we want to consider more complicated physics but our analytical capabilities aren't up to solving the more realistic problems, we need a perturbation theory. Since our knowledge of analytical geometry is almost intuitive we'll feel more comfortable if the perturbation theory is expressed in terms of the orbital element set $\underline{a} = (a,e,w,i,\Omega,T)$ or some function thereof. Hence, confining ourselves to a conservative disturbing force $+m\nabla R$ for the moment, we derive the perturbation equations equivalent to

$$m\ddot{\underline{r}} = -m\nabla U + m\nabla R$$

where U is the zero'th order potential. So set $\underline{r} = \underline{r}(\underline{a},t)$ and derive that

$$\nabla_{\underline{a}} \frac{\partial \underline{r}}{\partial t} \cdot \dot{\underline{a}} = \nabla R$$

Along the way I've imposed the condition of osculation on $\underline{r}$,

$$\dot{\underline{r}} = \partial \underline{r}/\partial t \quad \text{or} \quad \nabla_{\underline{a}} \underline{r} \cdot \dot{\underline{a}} = 0$$

so that both locations and velocities can be computed from the usual (eg. unperturbed) formulas. Finally we rearrange all of this via Lagrange brackets to read

$$\sum_{k=1}^{6} [a_\ell, a_k]\dot{a}_k = \partial R/\partial a_\ell \qquad (14)$$

where the Lagrange bracket $[a_\ell, a_k]$ is defined by

$$[a_\ell, a_k] = \left(\frac{\partial x}{\partial a_\ell}\frac{\partial \dot{x}}{\partial a_k} - \frac{\partial x}{\partial a_k}\frac{\partial \dot{x}}{\partial a_\ell}\right) + (x \to y) + (x \to z) \qquad (15)$$

We note that $d[a_\ell, a_k]/dt = 0$ so that the Lagrange brackets can be evaluated at any convenient place in the orbit. I choose M=0 (t=T). We further note that there are $\binom{6}{2} = 15$ independent brackets because a Lagrange bracket is antisymmetric and the most there could be is $6^2=36$. The last point of interest is to remember Eq. (12) for $\underline{r}$ and realize that there are three types of Lagrange brackets:

Type I: both a_ℓ, a_k are one of i, ω, Ω of which there are $\binom{3}{2} = 3$ such,

Type II: a_ℓ is one of i, ω, Ω but a_k is one of a,e,T (or M) of which there are $3 \cdot 3 = 9$ such and

Type III: both a_ℓ, a_k are one of a,e,T (or M) of which there are $\binom{3}{2} = 3$.

So far this is (formally) just like the two-body problem. In fact all of the brackets not involving a are the same as therein and the whole set of non-zero results is

$$[a,M] = -an(2-e^2) = -[M,a]$$

$$[a,\Omega] = -2bn\cos i = -[\Omega,a]$$

$$[a,\omega] = -2bn = -[\omega,a]$$

$$[e,M] = a^2 ne = -[M,e]$$

$$[e,\Omega] = \frac{a^2 ne\cos i}{(1-e^2)^{1/2}} = -[\Omega,e]$$

$$[e,\omega] = \frac{a^2 ne}{(1-e^2)^{1/2}} = -[\omega,e]$$

$$[i,\Omega] = abn\sin i = -[\Omega,i]$$

Alternatively $[T,e] = a^2n^2e$, $[T,a] = an^2e^2$.

The last step is to put these results into Eq. (14) and explicitly solve for $\dot{a}$. One finds

$$\frac{da}{dt} = \frac{1}{ane^2}\frac{\partial R}{\partial M} - \frac{(1-e^2)^{1/2}}{ane^2}\frac{\partial R}{\partial \omega}$$

$$\frac{de}{dt} = \frac{2(1-e^2)}{a^2ne^3}\frac{\partial R}{\partial M} - \frac{(2-e^2)(1-e^2)^{1/2}}{a^2ne^3}\frac{\partial R}{\partial \omega} \qquad (16)$$

$$\frac{d\omega}{dt} = \frac{(1-e^2)^{1/2}}{a^2ne^3}\left[(2-e^2)\frac{\partial R}{\partial e} + ae\frac{\partial R}{\partial a} - \frac{e^3\cot i}{1-e^2}\frac{\partial R}{\partial i}\right]$$

$$\frac{di}{dt} = \frac{\csc i}{na^2(1-e^2)^{1/2}}\left(\cos i\frac{\partial R}{\partial \omega} - \frac{\partial R}{\partial \Omega}\right)$$

$$\frac{d\Omega}{dt} = \frac{\csc i}{na^2(1-e^2)^{1/2}}\frac{\partial R}{\partial i}$$

$$\frac{dM}{dt} = \frac{-1}{nae^2}\frac{\partial R}{\partial a} - \frac{2(1-e^2)}{a^2ne^3}\frac{\partial R}{\partial e}$$

IV. THE ANISOTROPIC, THREE DIMENSIONAL, SIMPLE HARMONIC OSCILLATOR

Before delving into the main subject of this Section it seems prudent to me to review the mathematical basis of perturbation theory within celestial mechanics and its relatives in field theory. This will clearly reveal the difficulties which those of us who work in celestial mechanics face while demonstrating our collective blindness to the "right" coordinate system (only Vinti [1961] has overcome this but in a trivial way). Then the problem will be completely solved, in a variety of ways, in the right coordinate system.

A. The Mathematical Foundations of Perturbation Theory

Within classical mechanics we deal with equations of motion of the form

$$\underline{F} = m\underline{a}$$

If we set $\underline{v} = d\underline{r}/dt$, then an equivalent first order system is

$$\underline{v} = d\underline{r}/dt, \quad \underline{F} = md\underline{v}/dt$$

If $\underline{F}$ is conservative, $\underline{F} = -\nabla V$, then we have

$$\underline{v} = d\underline{r}/dt \ , \ -\nabla V = md\underline{v}/dt \tag{17}$$

Whenever V is such that we can't solve the problem analytically, we try to separate it into two parts. One part, represented by U, is such that Eqs. (17) are exactly soluble, while the other part (represented by -R in the standard sign convention) makes the problem intractable. The solution of

$$\underline{v} = d\underline{r}/dt, \quad -\nabla U = md\underline{v}/dt$$

has six arbitrary constants associated with it, say $\underline{a}$. The parametrization

of the solution, $\underline{r} = \underline{r}\ (\underline{a},t)$, is then used, as discussed above, to formulate $\underline{a}$. Since $|R|$ is small in some sense we can rewrite it as ϵR where ϵ is a small parameter. Hence we are led to study the solution of equations of the form (or systems of such equations)

$$dy/dx = \epsilon f(x,y)\ ,\quad y(x_0) = y_0 \tag{18}$$

The development below is independent of the dimensionality of y.

I know of one rigorous method of approximately solving Eq. (18)--Picard's method of successive substitutions. Glossing over the mathematical details (see Ince 1927) we construct the sequence $\{y_n(x)\}$,

$$y_n(x) = y_0 + \epsilon \int_{x_0}^{x} f[x', y_{n-1}(x')]dx',\ n = 1,2,\ldots$$

The sequence $\{y_n - y_{n-1}\}$ converges uniformly and absolutely to a function $y(x)$ which is continuous and satisfies Eq. (18). The solution of Eq. (18) is unique and stable with respect to volume perturbations (eg. small changes in f) and surface perturbations (eg. small changes in y_0). That's it in a nutshell.

Computationally carrying out the Picard process is enormously difficult in general. To my knowledge it has never been carried out beyond n=1 in celestial mechanics. The reasons for this are two: the orbital element set is an extraordinarily poor reference frame which exploits none of the symmetries of R (or U) and the interesting problems of celestial mechanics (atmospheric drag, oblateness perturbations, and third-body effects) are not simple. In an effort to deal with these difficulties other forms of

approximation have been tried. One is to try and obtain a solution to Eq. (18) in a power series in ϵ. For instance one would calculate y_1 correctly

$$Y_1(x) = y_1(x) = y_0 + \epsilon \int_{x_0}^{x} f(x_1, y_0) dx_1$$

but instead of computing y_2 from

$$y_2(x) = y_0 + \epsilon \int_{x_0}^{x} f[x_1, y_1(x_1)] dx_1$$

replace y_1 by its expression,

$$y_2(x) = y_0 + \epsilon \int_{x_0}^{x} f[x_1, y_0 + \epsilon \int_{x_0}^{x_1} f(x_2, y_0) dx_2] dx_1$$

and then expand f in a power series in ϵ ($\partial f(x,y) = \partial f(u,v)/\partial v|_{u=x, v=y}$)

$$y_2(x) \simeq y_0 + \epsilon \int_{x_0}^{x} \left\{ f(x_1, y_0) + \epsilon \partial f(x_1, y_0) \left[\int_{x_0}^{x_1} f(x_2, y_0) dx_2 \right] \right\} dx_1$$

$$Y_2(x) = y_0 + \epsilon \int_{x_0}^{x} f(x_1, y_0) dx_1 + \epsilon^2 \int_{x_0}^{x} \partial f(x_1, y_0) \left[\int_{x_0}^{x_1} f(x_2, y_0) dx_2 \right] dx_1$$

If one continues this policy to higher orders--expanding f at each step so that for Y_3

$$y_3(x) = y_0+\epsilon\int_{x_0}^{x} f[x_1,y_2(x_1)]dx_1$$

$$= y_0+\epsilon\int_{x_0}^{x} f\left\{x_1,y_0+\epsilon\int_{x_0}^{x_1} f[x_2,y_1(x_2)]dx_2\right\}dx_1$$

$$= y_0+\epsilon\int_{x_0}^{x} \left(f(x_1,y_0) +\epsilon\partial f(x_1,y_0)\left\{\int_{x_0}^{x_1} f[x_2,y_1(x_2)]dx_2\right\}\right)dx_1$$

$$= y_0+\epsilon\int_{x_0}^{x} f(x_1,y_0)dx_1+\epsilon^2\int_{x_0}^{x} \partial f(x_1,y_0)\left\{\int_{x_0}^{x_1} f[x_2,y_0+\epsilon\int_{x_0}^{x_2} f(x_3,y_0)dx_3]dx_2\right\}dx_1$$

$$\simeq y_0+\epsilon\int_{x_0}^{x} f(x_1,y_0)dx_1+\epsilon^2\int_{x_0}^{x} \partial f(x_1,y_0)\left(\int_{x_0}^{x_1}\left\{f(x_2,y_0)\right.\right.$$

$$\left.\left. + \epsilon\partial f(x_2,y_0)\left[\int_{x_0}^{x_2} f(x_3,y_0)dx_3\right]\right\}dx_2\right)dx_1$$

or

$$Y_3 = y_0+\epsilon\int_{x_0}^{x} f(x_1,y_0)dx_1+\epsilon^2\int_{x_0}^{x} \partial f(x_1,y_0)\left[\int_{x_0}^{x_1} f(x_2,y_0)dx_2\right]dx_1$$

$$+ \epsilon^3\int_{x_0}^{x} \partial f(x_1,y_0)\left\{\int_{x_0}^{x_1} \partial f(x_2,y_0)\left[\int_{x_0}^{x_2} f(x_3,y_0)dx_3\right]dx_2\right\}dx_1$$

$$= Y_2 + \epsilon^3 \text{ term}$$

or, in general,

$$Y_n(x) = Y_{n-1} + \epsilon \int_{x_0}^{x^n} \partial f(x_1, y_0) \ldots \left(\int_{x_0}^{x_{n-3}} \partial f(x_{n-2}, y_0) \left(\int_{x_0}^{x_{n-2}} \partial f(x_{n-1}, y_0) \right.\right.$$

$$\left.\left. \int_{x_0}^{x_{n-1}} f(x_n, y_0) dx_n \right) dx_{n-1} \right) \; dx_{n-2} \ldots dx_1$$

--then one can prove that the sequence $\{Y_n\}$ converges absolutely and uniformly to a continuous function of x that satisfies the initial condition. The series <u>does not</u> satisfy the differential equation. Hence it is useless.

If instead of substituting, expanding, substituting etc. one substitutes and then expands one gets a different result (beyond y_2 or Y_2). For instance return to y_3,

$$y_3(x) = y_0 + \epsilon \int_{x_0}^{x} f[x_1, y_2(x_1)] dx_1$$

$$= y_0 + \epsilon \int_{x_0}^{x} f\left\{x_1, y_0 + \epsilon \int_{x_0}^{x_1} f[x_2, y_1(x_2)] dx_2\right\} dx_1$$

$$= y_0 + \epsilon \int_{x_0}^{x} f\left\{x_1, y_0 + \epsilon \int_{x_0}^{x_1} f[x_2, y_0 + \epsilon \int_{x_0}^{x_2} f(x_3, y_0) dx_3] dx_2\right\} dx_1$$

Now expand from the inside outwards

$$Y_3(x) \simeq y_0 + \epsilon \int_{x_0}^{x} f\Big\{x_1, y_0 + \epsilon \int_{x_0}^{x_1} \Big[f(x_2,y_0) + \epsilon \partial f(x_2,y_0) \Big(\int_{x_0}^{x_2} f(x_3,y_0)dx_3\Big)\Big] dx_2 \Big\} dx_1$$

$$= y_0 + \epsilon \int_{x_0}^{x} f\Big\{x_1, y_0 + \epsilon \int_{x_0}^{x_1} f(x_2,y_0)dx_2 + \epsilon^2 \int_{x_0}^{x_1} \partial f(x_2,y_0) \Big(\int_{x_0}^{x_2} f(x_3,y_0)dx_3\Big) dx_2\Big\} dx_1$$

$$\simeq y_0 + \epsilon \int_{x_0}^{x} \Big\{f(x_1,y_0) + \Big[\epsilon \int_{x_0}^{x_1} f(x_2,y_0)dx_2$$

$$+ \epsilon^2 \int_{x_0}^{x_1} \partial f(x_2,y_0) \Big(\int_{x_0}^{x_2} f(x_3,y_0)dx_3\Big) dx_2\Big] \partial f(x_1,y_0)$$

$$+ \frac{\epsilon^2}{2}\Big[\int_{x_0}^{x_1} f(x_2,y_0)dx_2\Big]^2 \partial^2 f(x_1,y_0)\Big\} dx_1$$

through all terms of order ϵ^2 ($\partial^2 f(x,y) = \partial^2 f(u,v)/\partial v^2|_{u=x,v=y}$). This is not equal to y_3 or Y_3 but the technique represents a perfectly acceptable approximation scheme (on the surface). Note too that the Y_n expansion only required that $\partial f/\partial y$ exists while this will require that all partial derivatives of f with respect to y exist and be continuous.

The method of averaging is yet another way around the difficulty which is not rigorous either (so far as I'm aware). Indeed in the next subsection the use of it on the harmonic oscillator will illustrate it's shortcomings.

Now what does perturbation theory for a scalar field look like? Suppose that we are dealing with

$$\nabla^2\psi + (k^2 - U_0 - \lambda U)\psi = 0$$

where the unperturbed problem is

$$\nabla^2\psi + (k^2 - U_0)\psi = 0$$

Let the solutions be the complete set of orthonormal eigenfunctions u_n corresponding to the eigenvalue k_n^2 (degeneracy is an annoyance not dealt with here). We found these eigenfunctions by exploiting the symmetries of U_0, not by restricting ourselves to the same coordinate system for all perturbations. The approximations to the perturbed eigenfunctions and eigenvalues are given by

$$\psi^{(1)} = u_n + \lambda \sum_{p \neq n} \frac{U_{pn}}{k_n^2 - k_p^2} u_p$$

$$(k^2)^{(1)} = k_n^2 + \lambda U_{nn}$$

$$\psi^{(2)} = u_n + \lambda \sum_{p \neq n} \frac{U_{pn} u_p}{(k_n^2 + \lambda U_{nm} - k_p^2)} + \lambda^2 \sum_{pq \neq n} \frac{U_{pq} U_{qn} u_p}{(k_n^2 - k_p^2)(k_n^2 - k_q^2)}$$

$$(k^2)^{(2)} = k_n^2 + \lambda U_{nn} + \lambda^2 \sum_{p \neq n} \frac{U_{np} U_{pn}}{k_n^2 - k_p^2}$$

and for the a'th order

$$\psi^{(a)} = u_n + \lambda \sum_{p \neq n} \frac{U_{pn} u_p}{[(k^2)^{(a-1)} - k_p^2]} + \ldots$$

$$+ \lambda^a \sum_{pq \ldots \neq n} \frac{U_{pq} \ldots U_{zn} u_p}{(k_n^2 - k_p^2)(k_n^2 - k_q^2) \ldots (k_n^2 - k_z^2)}$$

$$(k^2)^{(a)} = k_n^2 + \lambda U_{nn} + \lambda^2 \sum_{p \neq n} \frac{U_{np}U_{pn}}{[(k^2)^{(a-2)} - k_p^2]} + \cdots$$

$$+ \lambda^a \sum_{pq\ldots \neq n} \frac{U_{np}U_{pq}\ldots U_{zn}}{(k_n^2 - k_p^2)(k_n^2 - k_q^2)\ldots(k_n^2 - k_z^2)}$$

I've listed these formulas directly from Morse and Feshbach (1953) which the interested reader should consult. In general these series converge. Note that here the work consists of calculating the matrix elements of U and summing the series whereas in celestial mechanics the analytical work never ends.

B. The Method of Averaging

Kozai (1959) exploits the method of averaging in order to solve the main problem of artificial satellite theory. This is defined as to solve for the motion of a particle under a potential which includes, in addition to the usual 1/r term, the first three oblateness terms assuming axial symmetry--namely the J_2,J_3 and J_4 terms. Various developments of the method of averaging (Lorell, Anderson, and Lass 1964; Lorell and Liu, 1971; McClain 1977) all refer back to Kryloff and Bogoliuboff (1947) and Bugoliubov and Mitropolsky (1961) for a basis. If this method has a rigorous foundation I can't find it. As an example consider the following (pgs. 39-41 of the last mentioned reference):

> We now go over to the discussion of a method for developing the first instance, asymptotic approximations, for the case of oscillations defined by a differential equation of the form:
>
> $$\frac{d^2x}{dt^2} + \omega^2 x = \varepsilon f\left(x, \frac{dx}{dt}\right), \tag{1.1}$$
>
> where ε is a small positive parameter. We can arrive at the correct formulation of this method if we start from the physical concepts defining the character of the oscillatory process.
>
> When perturbation is absent i.e. when $\varepsilon = 0$, the oscillations will, evidently be purely harmonic,
>
> $$x = a \cos \psi$$
>
> with a constant amplitude and a uniformly rotating phase angle:

$$\frac{da}{dt} = 0, \quad \frac{d\psi}{dt} = \omega \quad (\psi=\omega t+\theta),$$

(the amplitude a and the phase θ of the oscillations will be constants over time, depending on the initial conditions).

The existence of non-linear perturbation ($\varepsilon \neq 0$) results in the appearance of overtones in the solution of equation (1.1), a factor that establishes dependence between the instantaneous frequence $\frac{d\psi}{dt}$ and the amplitude, and finally gives rise to a systematic increase or decrease in the amplitude of the oscillations, depending upon whether the energy is expelled or absorbed by the perturbing forces. All these effects are obviously in the limiting case ($\varepsilon=0$).

With all this in view we shall seek a general solution of equation (1.1) in the form,

$$x = a\cos\psi + \varepsilon u_1(a,\psi) + \varepsilon^2 u_2(a,\psi) + \varepsilon^3 u_3(a,\psi) + \ldots \tag{1.2}$$

Here $u_1(a,\psi)$, $u_2(a,\psi)\ldots$ are periodic functions of the angle ψ with a period 2π and the quantities a,ψ are functions of time defined by the differential equations:

$$\left.\begin{aligned} \frac{da}{dt} &= \varepsilon A_1(a) + \varepsilon^2 A_2(a) + \ldots, \\ \frac{d\psi}{dt} &= \omega + \varepsilon B_1(a) + \varepsilon^2 B_2(a) + \ldots \end{aligned}\right\} \tag{1.3}$$

We are to choose suitable expressions for the functions $u_1(a,\psi)$, $u_2(a,\psi)\ldots$, $A_1(a)$, $B_1(a)$, $A_2(a)$, $B_2(a),\ldots$ in such a way that the equation (1.2), after replacing a and ψ, by the functions defined in equations (1.3), would serve as a solution of equation (1.1).

As soon as this problem is solved and explicit expressions for the coefficients of expansions occuring in the right-hand-sides of equations (1.2) and (1.3) are obtained, the problem of integrating equation (1.1) is reduced to that of integrating equations (1.3) which have separable variables, thus making the investigation possible with the help of well known elementary methods.

. . . . . .

We note that in the case represented by equation (1.1) we might establish the convergence of expansions (1.2), (1.3) under very general conditions for the function $f\left(x,\frac{dx}{dt}\right)$.

However, since in future we will have to deal with cases in which similar expansions apparently diverge, we will not tie up the development of our method of the construction of asymptotic approximations with any proof of convergence.

I find it difficult to be so cavalier. What we're supposed to do is take R

$$R = -\varepsilon n^2 z^2/2 = -(\varepsilon n^2 a^2/2)\sin^2 i[\sin\omega\cos M+(1-e^2)^{1/2}\cos\omega\sin M]^2$$

and form <R>,

$$\langle R\rangle = \frac{1}{2\pi}\int_0^{2\pi} R\,dM$$

$$= \frac{-\varepsilon n^2 a^2}{4\pi}\sin^2 i[\pi\sin^2\omega + 2(1-e^2)^{1/2}\sin\omega\cos \cdot 0$$

$$+ \pi(1-e^2)\cos^2\omega] \qquad (19)$$

$$= \frac{-\varepsilon n^2 a^2}{4}\sin^2 i\,[\sin^2\omega+(1-e^2)\cos^2\omega]$$

If a term of <R> depends upon an element that has a secular term, and the dependence upon this element is period (eg the $\sin^2\omega$,$\cos^2\omega$ terms), then we are to say that these are long periodic terms. If the dependence is not periodic, then these are secular terms. R-<R> represents the short periodic terms. Does <R> have a secular term? Yes if you say rewrite it as

$$\langle R\rangle = \frac{-\varepsilon n^2 a^2 \sin^2 i}{4}(1-e\cos^2\omega)$$

for then

$$R_{sec} = -\varepsilon n^2 a^2 \sin^2 i/4$$

$$R_{\ell p} = \varepsilon n^2 a^2 e^2 \sin^2 i\cos^2\omega/4$$

$$R_{sp} = R - R_{sec} - R_{\ell p}$$

No if you don't. Apparently the method of averaging and the assignment of terms to various parts of R depends on one's mathematical sophistication.

Ignoring the above dilemnas let's turn to the solution for the orbital element sets. We're supposed to obtain the short period changes in a, e, etc. by using R_{sp} in Eqs. (16). This is fine for first order effects except for the $\partial R/\partial M$ terms because the operations $\partial/\partial M$ and $\int dt$ don't commute. To evaluate the first order secular terms one proceeds in a similar fashion. In both instances one regards the orbital elements appearing on the right hand sides of Eqs. (16) as constants. Again this is fine for first order, short term (times $< P/\varepsilon$) results. Finally one is supposed to obtain the long period changes by an exactly analogous procedure. Since ω varies by 2π on a time scale of P/ε, the periodicity isn't apparent until a duration of order P/ε^2, and regarding a,e,i,etc. to be constants for this long is of dubious validity.

I have not repeated these computations here because they're apparently not rigorous, they don't illuminate the nature of the motion (short term, secular, or long term) and they can't be carried out in an unambiguous fashion (if I can't have rigor, I at least demand explicit instructions). Had I produced the results and interpreted them, I would've drawn some interesting, and false, conclusions regarding the motion (remember that this is an exactly soluble problem whose solution is periodic). Hence I intend to turn to a removal of part of the problem associated with perturbation theory in celestial mechanics--the orbital element set is a terrible basis for a coordinate system.

C. A New Representation

The above "solution" of the perturbation equations leaves much to be desired--like an answer. The problem is the basis of the representation $\underline{a} = (a,e,w,i,\Omega,T)$. Go back to the original, unperturbed equations of motion ($U = kr^2/2$),

$$m\ddot{\underline{r}} = \underline{F} = -\nabla U = -k\underline{r}$$

The general solution is

$$\underline{r} = \underline{A}\cos nt + \underline{B}\sin nt \qquad (20)$$

where, cf. Eq. (5), $n^2=k/m$. If we add a perturbing force $+m\nabla R$ the values of $\underline{A}$ and $\underline{B}$ will no longer be constant. Let's develop a perturbation theory for $\underline{a} = (\underline{A},\underline{B})$ now. Well the work is all done and is given by Eqs. (14,15). Once *again there are only* 15 *independent* Lagrange brackets to evaluate since 21 of them obviously vanish. These are of three types too:

Type I: both a_ℓ, a_k are one of the elements of $\underline{A}$ of which there are $\binom{3}{2} = 3$ such,

Type II: a_ℓ is one of the elements of $\underline{A}$ but a_k is one of the elements of $\underline{B}$ of which there are $3\cdot 3=9$ such, and

Type III: both a_ℓ, a_k are one of the elements of $\underline{B}$ of which there are $\binom{3}{2} = 3$.

All Type I and III brackets vanish as do the "non-diagonal" Type II's. The only non-zero ones are

$$[A_x,B_x] = [A_y,B_y] = [A_z,B_z] = n$$

and their negatives. The simplicity of this result already makes one suspect that this is a very nice basis.

The perturbed equations of motion are

$$m\ddot{\underline{r}} = -k\underline{r} - (k-k)z(0,0,1)$$

Let $k = k(1+\epsilon)$ with $|\epsilon|<<1$ as before. Then $R = -\epsilon n^2 z^2/2$ or

$$R = \frac{-\epsilon n^2}{2}(A_z \cos nt + B_z \sin nt)^2$$

The variation of parameters equations are just

$$\dot{A}_x = \dot{A}_y = \dot{B}_x = \dot{B}_y = 0$$

$$\dot{A}_z = \epsilon n(A_z \cos nt + B_z \sin nt)\sin nt \qquad (21)$$

$$\dot{B}_z = -\epsilon n(A_z \cos nt + B_z \sin nt)\cos nt$$

Much simpler it would be difficult to imagine. Moreover because these equations are first order, linear, ordinary differential equations with polynomial or exponential coefficients Picard's method of successive approximations can be carried out, explicitly, to all orders.

Let the zero'th order approximation be denoted by $A_z(0), B_z(0)$. These are simply related to the initial values for z and $\dot{z}$ in the unperturbed case,

$$z(0) = A_z(0),\ \dot{z}(0) = nB_z(0)$$

Using primes to indicate the order in Picard's method we have, successively

$$\dot{A}'_z = \varepsilon n[A_z(0)\cos nt + B_z(0)\sin nt]\sin nt$$

$$\dot{B}'_z = -\varepsilon n[A_z(0)\cos nt + B_z(0)\sin nt]\cos nt$$

$$A'_z = A_z(0) + (\varepsilon/2)A_z(0)\sin^2 nt + (\varepsilon/2)B_z(0)(nt-\sin nt\cos nt)$$

$$B'_z = B_z(0) - (\varepsilon/2)B_z(0)\sin^2 nt - (\varepsilon/2)A_z(0)(nt+\sin nt\cos nt)$$

$$\dot{A}''_z = \varepsilon n(A'_z\cos nt+B'_z\sin nt)\sin nt$$

$$\dot{B}''_z = -\varepsilon n(A'_z\cos nt+B'_z\sin nt)\cos nt$$

$$A''_z = A'_z - (\varepsilon^2/8)A_z(0)[n^2t^2+(\sin nt-2nt\cos nt)\sin nt] +$$

$$(\varepsilon^2/16)B_z(0)(3\sin 2nt-4nt-2nt\cos 2nt)$$

$$B''_z = B'_z + (\varepsilon^2/8)B_z(0)[-n^2t^2+(3\sin nt-2nt\cos nt)\sin nt]$$

$$+ (\varepsilon^2/16)A_z(0)(\sin 2nt-2nt\cos 2nt)$$

$$\dot{A}'''_z = n(A''_z\cos nt+B''_z\sin nt)\sin nt$$

$$\dot{B}'''_z = -n(A''_z\cos nt+B''_z\sin nt)\cos nt$$

etc. We see short period terms, secular terms, and mixed terms in the above. Long period terms will never appear though.

The advantages of the $\underline{A},\underline{B}$ representation over the orbital element set representation should be clear now. Since Picard's method can be rigorously

applied to all orders all we need to do is sum the series to get closed form results for A_z and B_z. Of course for this problem a much simpler course is available to us for there's an even better basis than the $\underline{A},\underline{B}$ one.

Define α and β via

$$\alpha = A_z\cos nt + B_z \sin nt, \quad \beta = -A_z \sin nt + B_z \cos nt \tag{22}$$

They satisfy

$$\dot{\alpha} = n\beta, \quad \dot{\beta} = -n\alpha(1+\epsilon)$$

or

$$\ddot{\alpha} = -n^2(1+\epsilon)\alpha$$

The solution to this can be written down by inspection,

$$\alpha = C\cos \nu t + D \sin \nu t$$

where C and D are the arbitrary constants of integration and

$$\nu^2 = n^2(1+\epsilon)$$

Also $\beta = \dot{\alpha}/n = (\nu/n)(-C\sin \nu t + D\cos \nu t)$. We can recover A_z and B_z from Eqs. (22),

$$A_z = \alpha\cos nt - \beta \sin nt, \quad B_z = \alpha \sin nt + \beta \cos nt$$

Or

$$A_z = C[\cos nt\cos\nu t+(\nu/n)\sin nt\sin\nu t]$$

$$+ D[\cos nt\sin\nu t-(\nu/n)\sin nt\cos\nu t] \qquad (23)$$

$$B_z = C[\sin nt\cos\nu t-(\nu/n)\cos nt\sin\nu t]$$

$$+ D[\sin nt\sin\nu t+(\nu/n)\cos nt\cos\nu t]$$

We can also identify C and D as

$$C = A_z(0), \quad D = (n/\nu)B_z(0) = B_z(0)/(1+\varepsilon)^{1/2}$$

Complete success--the exact, analytical solution of the perturbation equations.

It's appropriate now to reflect on our accomplishments. First of all we found a basis for the development of a perturbation theory (a set of arbitrary constants) which yield a much simpler set of variation of parameter equations that we would've expected. Secondly we can solve this by actually implementing Picard's method of successive approximations. Hence, we now know that a solution exists, it's continuous, it's unique, etc. Moreover, <u>for this problem</u>, the n'th order successive approximation in Picard's method is of order ε^n. This coincidence is due to the linearity of Eqs. (21). See the next section for a counterexample to this result as a general proposition. Thirdly the full set of variation of parameters equations are exactly soluble in closed form. This reflects the fact that the original problem was analytically tractable (since $\alpha \equiv z$, $\beta \equiv \dot{z}/n$) <u>and</u> the choice of the correct representation. Finally should one choose to expand the direct solution for A_z, B_z in Eqs. (23) in a Taylor series in ε one will recover, order by order, the successive steps of the Picard scheme. As I wrote above, a complete success.

One last point. An argument used in favor of the orbital element basis for perturbation theory was that doing so made the interpretation of the results easier. I claim that the form of Eqs. (23) makes it crystal clear that the three dimensional anisotropic, simple harmonic oscillator is space filling.

V. THE ANHARMONIC, ONE DIMENSIONAL, HARMONIC OSCILLATOR

In this Section I want to discuss anharmonic perturbations and since the harmonic oscillator problem separates in rectangular coordinates, one dimension is sufficient. The unperturbed equations of motion are

$$m\ddot{x} = -kx \quad , \; n^2 = k/m$$

whose general solution is

$$x = A\cos nt + B\sin nt$$

The perturbed equations of motion are

$$m\ddot{x} = -kx - \eta kx^3$$

so $R = -\eta n^2 x^4/4$ (I could've chosen a quadratic form for the anharmonic term but this choice simplifies the analysis somewhat). The basis for the perturbation theory is $\underline{a} = (A,B)$. The only non-zero Lagrange bracket is $[A,B] = -[B,A] = n$. The variation of parameter equations are

$$\dot{A} = -(1/n)\partial R/\partial B, \quad \dot{B} = (1/n)\partial R/\partial A$$

Since $R = -\eta n^2 x^4/4$,

$$R = (-\eta n^2/4)\,(A\cos nt + B\sin nt)^4$$

or

$$\dot{A} = \eta n \sin nt\,(A\cos nt + B\sin nt)^3$$

$$\dot{B} = -\eta n \cos nt\,(A\cos nt + B\sin nt)^3$$

Let A', B' be the first order set of results from Picard's method, viz.

$$A'(t)=A(0)+(\eta/32)[8A^3(0)(1-\cos^4 nt)+3A^2(0)B(0)(4nt-\sin 4nt)$$
$$+24A(0)B^2(0)\sin^4 nt+B^3(0)(12nt-8\sin 2nt+\sin 4nt)]$$
$$B'(t)=B(0)-(\eta/32)[A^3(0)(12nt+8\sin 2nt+\sin 4nt)+24A^2(0)B(0)(1-\cos^4 nt)$$
$$+3A(0)B^2(0)(4nt-\sin 4nt)+8B^3(0)\sin^4 nt]$$

Clearly the next approximation, obtained from

$$\dot{A}'' = \eta n \sin nt(A'\cos nt+B'\sin nt)^3$$
$$\dot{B}'' = -\eta n \cos nt(A'\cos nt+B'\sin nt)^3$$

results in powers of η of order 4 not just order 2. Just as clearly an anharmonic perturbation in the equation of motion of the form $-\eta kx^{p-1}$ results in A",B" containing terms of the form η^p. I hope that this result clearly illustrates the difference between the order (in the sequence of successive approximations) of the solution and the functional form or highest power present of a small parameter.

The perturbation equations can be solved analytically. To see how define α and β via

$$\alpha = A\cos nt+B\sin nt$$
$$\beta = -A\sin nt+B\cos nt$$

Then, because of the condition of osculation,

$$\dot{\alpha} = n\beta, \quad \dot{\beta} = -n\alpha - \eta n\alpha^3$$

or

$$\dot{\alpha} = -n^2\alpha - \eta n^2\alpha^3$$

A first integral is

$$\dot{\alpha}^2 = n^2(L^2-\alpha^2) + (\eta n^2/2)(L^4-\alpha^4)$$

if the initial conditions are $x(0)=L$, $\dot{x}(0)=0$. Other initial conditions do not yield more transparent solutions or allow me to make additional analytical points.

The solution for α is

$$\alpha = Lcn[nt(1+\eta L^2)^{1/2},k]$$

$$k^2 = L^2/(2L^2+2/\eta) \ , \ k\varepsilon[0,1]$$

where cn is the cosine-amplitude Jacobian elliptic function. Since $\beta = \dot{\alpha}/n$,

$$\beta = -L(1+\eta L^2)^{1/2}sn[nt(1+\eta L^2)^{1/2} \ , \ k]dn[nt(1+\eta L^2)^{1/2}; \ k]$$

where sn is the sine-amplitude Jacobian elliptic function and dn is the delta-amplitude Jacobian elliptic function. The solutions for A and B are $[m=nt(1+\eta L^2)^{1/2}]$

$$A/L = cosntcn(m,k)+(1+\eta L^2)^{1/2}sinntsn(m,k)dn(m,k)$$

$$B/L = sinntcn(m,k)-(1+\eta L^2)^{1/2}cosntsn(m,k),dn(m,k)$$

As $\eta \to 0$

$$k^2 \to \eta L^2/2$$

$$sn(m,k) \to sinm-(k/2)^2cosm(m-sinmcosm)$$

$$cn(m,k) \to cosm+(k/2)^2sinm(m-sinmcosm)$$

$$dn(m,k) \to 1 -(k^2/2)^2sin^2m$$

By some laborious algebra one can show that the first order Taylor series expansions of A and B exactly match the expressions for A' and B'[A(0) = L,B = 0]. Clearly the statement is not true for the second order Taylor series since A" and B" contain terms proportional to η^4. I don't know if the fourth order Taylor series expansions of A and B would match A" and B"--I have only a finite amount of patience (I would bet on it though).

Lastly let me mention that the original equations of motion are exactly soluble too. For $x(0)=L$, $\dot{x}(0)=0$ the solution is

$$x = Lcn(m,k)$$
$$\rightarrow Lcosnt \text{ as } \eta\rightarrow 0$$

while for $x(0)=0$, $\dot{x}(0)>0$ it is

$$x = (L^2+2/\eta)^{1/2}\, sn(m,k)/dn(m,k)$$

where max(x)=L. As $\eta\rightarrow 0$ this approaches Lsinnt. The general solution is not a linear combination of these two (it's anharmonic) nor is it worth displaying.

VI. CONCLUDING REMARKS

Another way of writing the general solution to the harmonic oscillator (one dimensional for the moment) is

$$x = C\cos(nt+\phi)$$

The Lagrange bracket $[C,\phi]=-nC$ and the perturbation equations are

$$-nC\dot{\phi} = \partial R/\partial C, \quad nC\dot{C} = \partial R/\partial \phi$$

But R is time dependent for either anharmonic or anisotropic perturbations. This obviously complicates the solution and I haven't explored this alternative in depth. Another thing I haven't done is derive the analog of Gauss's equations for non-conservative perturbations. As an example if there is a drag term,

$$m\ddot{x} = -kx - 2m\gamma\dot{x}$$

then the general solution can be written as (I've assumed that γ is small compared to n, $n^2 = k/m$ still)

$$x = (A\cos\nu t + B\sin\nu t)e^{-\gamma t}$$

where $\nu^2 = n^2-\gamma^2$. It would be interesting to pursue this problem and see where else the usual form of perturbation theory in celestial mechanics let's us down.

An even more interesting generalization would be the simple harmonic oscillator subject to drag and an external, periodic force. The equation of motion would then be (say)

$$m\ddot{x} = -kx - 2m\gamma\dot{x} + F\cos(\omega t+\theta)$$

As you'll remember the steady state solution involves a resonance. The phase of x is offset from that of the external force by ψ

$$\tan\psi = \frac{2\gamma\omega}{n^2-\omega^2}$$

and the amplitude of x is $(F/m)/[(n^2-\omega^2)^2 + 4\gamma^2\omega^2]^{1/2}$. When the atmospheric drag is small, there is resonance at $\omega = n-\gamma^2/n$. I wonder what classical perturbation theory would do with this.

References

Arnold, V. I., Mathematical Methods of Classical Mechanics, Springer-Verlag, Berlin, 1978.

Bugoliubov, N. N. and Mitropolsky, Y. A., Asymptotic Methods in the Theory of Nonlinear Oscillations, Hindustan Publ. Corp., New Deli, 1961.

Goldstein, H., Classical Mechanics, Addison-Wesley Publ. Co., Inc., Reading, MA, 1982.

Kozai, Y., Astron. J. 64, 367, 1959.

Kryloff, N. M., Bogoliuboff, N. N., Introduction to Non-linear Mechanics, Publ. Acad. Sci. Ukr. SSR, 1937.

Lorell, J., Anderson, J. D., and Lass, H., J. P. L. TR. No. 32-482, 1964.

Lorell, J. and Liu, A., J. P. L. TR. No. 32-1513, 1971.

McClain, W. D., C. S. C. TR-77/6010, 1977.

Plummer, H. C., An Introductory Treatise on Dynamical Astronomy, Cambridge Univ. Press, London, 1918.

Vinti, J. P. J. Res. Nat'l. Bur. Stds. 65B, 169, 1961.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-83-023	2. GOVT ACCESSION NO. AD-A130213	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* On Perturbation Theory		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER Technical Report 650
7. AUTHOR(s) Laurence G. Taff		8. CONTRACT OR GRANT NUMBER(s) F19628-80-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lincoln Laboratory, M.I.T. P.O. Box 73 Lexington, MA 02173-0073		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Project Nos. 2698/2295 Program Element Nos. 63428F and 12424F
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Systems Command, USAF Andrews AFB Washington, DC 20331		12. REPORT DATE 1 June 1983 13. NUMBER OF PAGES 50
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)* Electronic Systems Division Hanscom AFB, MA 01731		15. SECURITY CLASS. *(of this report)* Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

perturbation theory	celestial mechanics	field theories
trajectory equations	field of averaging	series expansions

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Perturbation theory in celestial mechanics has a long, rich history of failure stretching back to Newton. I believe that the causes of this are two-fold. One problem is the difficulty of dealing with the mathematical structure used in celestial mechanics to express perturbation theory as opposed to the constructs used in field theories (eg. trajectory equations vs. linear second order partial differential equations). The second flows directly from this and relates to the misapplications of certain mathematical techniques (averaging, series expansions) within the context of perturbation theory.

This report clearly illustrates the nature of these difficulties utilizing a complex (but exactly soluble) physical model intimately tied to the two-body problem.

DD FORM 1 Jan 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*


FILM